DOCUMENT RESUME

ED 227 174 TM 830 203

AUTHOR Kaiser, Javaid
TITLE The Differential Predictive Validity of the GRE Aptitude Test for Foreign Students.
PUB DATE 24 Feb 83
NOTE 19p.; Presented at the Annual Meeting of the Eastern Educational Research Association (Baltimore, MD, February 24-27, 1983).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; *Aptitude Tests; *College Entrance Examinations; English (Second Language); *Foreign Students; Grade Point Average; Graduate Study; *Predictive Validity; Quantitative Tests; School Registration; Scores; Sex; *Test Bias; Testing Problems; Test Norms; Verbal Tests
IDENTIFIERS *Graduate Record Examinations; University of Kansas

ABSTRACT

The predictive validity of the Graduate Record Examinations (GRE) aptitude test has been investigated for foreign students at the University of Kansas. A total of 148 students whose first language was other than English were selected for this study. Verbal and quantitative scores on the GRE, field of study, sex, and year of initial enrollment were used as predictors. The grade point average (GPA) in the graduate school was considered as the criterion. The results revealed that foreign students scored significantly lower than American students on the GRE scores and the GPA. Poor correlation between the GRE scores and the criterion suggested that the GRE aptitude test is not the most appropriate way to predict academic performance of foreign students. During multiple regression, field of study and verbal scores on the GRE were found to be significant predictors of the GPA. High within-group variability among foreign students suggested the study of subgroups formed on the basis of language, culture, or the geographic location. The use of separate norms for foreign students was also found desirable. (Author/PN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED227174

The Differential Predictive Validity of the GRE Aptitude Test for Foreign Students

Javaid Kaiser

University of Kansas

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

J Kaiser

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Presented at the meeting of Eastern Educational Research Association at Baltimore, Maryland.

February 24, 1983.

TM 830 203

Abstract

The predictive validity of the GRE Aptitude test has been investigated for foreign students at the University of Kansas. A total of 148 students whose first language was other than English were selected for this study. Verbal and quantitative scores on the GRE, field of study, sex, and year of initial enrollment were used as predictors. The GPA in the graduate school was considered as the criterion. The results revealed that foreign students scored significantly lower than American students on the GRE scores and the GPA. Poor correlation between the GRE scores and the criterion suggested that the GRE Aptitude test is not the most appropriate way to predict academic performance of foreign students. During multiple regression, field of study and verbal scores on the GRE were found significant predictors of the GPA. The study suggested to investigate subgroups of foreign students formed on the basis of language, culture, or the geographic location. The use of separate norms for foreign students was also felt desirable.

The Differential Predictive Validity of the GRE Aptitude Test for Foreign students

The predictive validity of the GRE aptitude test for minority groups has long been questioned. Concern about possible test bias has resulted in scores of validation studies during the last two decades (Lannholm, 1960, 1972). The results of these studies include recommendations (1) to establish local norms and (2) to establish separate norms for the ethnic groups (Willingham, 1976). Foreign students who constitute about 6% of the graduate student population in United States were neglected in this validation process. Foreign students are required to take the the GRE Aptitude test. Their scores are used in a variety of selection decesions that affect their career and future plans. They represent a variety of cultures, languages, and educational systems and as a group are more vulnerable to test bias than American ethnic groups who have many things in common with the majority group.

The GRE Board became aware of this possible test bias a long time ago and initiated a study of 637 foreign students. The results revealed that foreign students scored lower on the GRE than their American counterparts. (Pitcher & Harvey, 1963). Shay and Tseng (1978) found similar results in a study of West Virginia University students. In another study, Harvey and Lannholm (1961) found no significant increase in the total scores of foreign students on the GRE aptitude test even after one year of graduate work at American universities. While a high discrepancy in the performance of American and foreign students on the verbal section of the GRE can be attributed to cultural bias, the discrepancy on

the GRE quantitative section stems, in part, from the same cause. The Measurement specialists have always emphasized the need to make the language of the test a neutral vehicle which is not true for the GRE aptitude test when given to foreign students. Angelis (1977) ,however, found little effect of increased language efficiency on the GRE scores for non-natives but considered the continious use of the GRE inappropriate for foreign students.

Conrad, Trisman, and Miller (1977) conducted a study of 131 minority subjects including foreign students. The results of this study were compared in terms of median r values with the corresponding median values computed from GRE validation studies done on American students for the period 1952 to 1972. The conclusion drawn by the authors that the GRE Aptitude test predicts graduate performance of natives and non-natives equally well is however, inappropriate because of the small sample sizes of subgroups on which these comparisons were made. In some groups this number was as small as 9.

Purpose of the study

The establishment of separate norms for the minority group is one of the key recommendations of the Educational Testing Service to institutions that use their tests. Most of the schools that attract a large number of foreign students lack normative data on this subgroup. In the abscence of norms, foreign students are either directly compared with their American counterparts or are subjected to a judgemental procedure for selection. The present study was designed to investigate the predictive validity of the GRE aptitude test for foreign students

and to help college personnel in establishing differential selection criteria for natives and non-natives.

Procedure

All the foreign students enrolled during the last five years at the University of Kansas for whom GRE scores were available, were the subjects in this study. The sample included 148 students. Information about the field of study (Major), the year of initial enrollment (Year), sex, GPA in the graduate school (GPA), and scores on the verbal (GRE-V) and quantitative (GRE-Q) subscales of the GRE, were collected on each student. Nominal variables like major was categorized into (1) physical sciences and (2) social sciences. Sex had two categories: (1) male and (2) female. Year had five levels: (1) 1973 or earlier, (2) 1974, (3) 1975, (4) 1976, and (5) 1977 or later. A composite GRE score (GRE-T) was also computed for each student by adding the verbal and quantitative scores together. GPA was used as the criterion while the rest of the variables were treated as predictors.

Two groups, one of 316,693 American students and another one of 21,881 foreign students, tested during October, 1974 and June, 1976 on the GRE Aptitude test, were treated as national samples reppresenting populations of American and foreign students, respectively. The statistic reported on these groups by Conrad et al. (1977) was considered as national norms. Means and standard deviations were computed on the GRE scores and the GPA. The statistics obtained on foreign students was compared with the national norms of American and foreign students as well as with the statistics reported by Kaiser (1982) on 407

American students. The data from Kaiser (1982) was also pooled with the data of this study to test the significance of mean differences on the GRE scores and the GPA between American and foreign students. A t-test was performed to determine the statistical significance. The test of homogeneity of variance was also conducted prior to the t-test to determine whether a pooled variance or a separate variance estimate should be used in the computation of t values.

The degree of relationship between the GRE scores and the GPA was determined by computing zero order correlation coefficients. The strength of relationship between the predictors and the criterion was investigated across groups of American and foreign students by comparing the obtained r values with the corresponding median r values, computed from validation studies completed prior to 1972 on American students. These median r values were reported by Conrad et al. (1977). The obtained correlation coefficients were also compared with the respective r values reported by Kaiser (1982) on the local reference group of American students.

Stepwise multiple regression was performed on the data to identify the best composite of multiple predictors for foreign students. Nominal variables were dummy coded into (n-1) independent vectors before their inclusion into the regression equation (Kerlinger & Pedhazer, 1973). For major, the subgroup of social sciences was used as reference group. Females were the reference point for sex. Similarly, 1973 or earlier was the reference group for year of enrollment. The interactions among major, sex, and year of enrollment were included in the analysis by creating dummy coded orthognal vectors.

The analysis was performed in two parts. In the first part, GRE-V and GRE-Q were included as separate predictors in the regression equation. In the second part, the total scores on the GRE were used as a single predictor instead of using verbal and quantitative scores, separately. Major, sex, year of enrollment, and their interactions were included in both the analysis. The stability of obtained R values was determined by computing their shrunken values (R) at each step, during the development of regression equation.

Results

Table 1 shows the means and standard deviations for the GRE-V, GRE-Q, GRE-T, and GPA on foreign students and local group of American students. The performance of national samples of American and foreign students on GRE-V and GRE-Q was also reported in this table. The scores of foreign students on verbal and quantitative subscales of the GRE were lower than the national group of foreign students. The statistics on GRE-T and GPA could not be compared due to the abscence of this information on national groups. Foreign students also scored lower than their local American counterparts on all the predictors and the GPA. Standard deviations were higher for foreign students than local American students on all the variables except GRE-V. This findings was consistent with the results obtained at national level where national sample of foreign students scored lower means and higher standard deviations on the GRE compared to the national sample of American students.

==========

Insert Table 1 Here

==========

Means of foreign students on the GRE scores and the GPA were tested against means of local American students for statistical significance The t-test was applied and the results were recorded in Table 2. The test of homogeneity of variance revealed significant differences in the population variances of GRE-Q and the GPA ($p < .01$). Separate variance estimates were therefore, used to compute t statistic for GRE-Q and GPA. The t values for GRE-V and GRE-T were based on pooled variance estimates. The t statistic indicated that the scores of foreign students were significantly lower than the scores of American students on all the predictors and the criterion except GRE-Q ($p < .01$). All the t values were based on 555 cases out of which 148 were foreign students.

==========

Insert Table 2 Here

==========

The intercorrelation among the predictors and the criterion were computed to determine the best single predictor of GPA. The results were recorded below the diagnol in Table 3. Correlation coefficients on the local group of American students, as reported by Kaiser (1982), were entered above the diagnol in the same table. The last two rows in the table show the median r values reported by Conrad et al. (1977) and the number of coefficients used in computing such values.

The results revealed that the GRE-V was the only significant predictor of the criterion for foreign students ($p < .05$). The GRE-Q and GRE-T were not significantly correlated with the GPA ($p > .05$). It was also observed that r values between the GRE scores and the GPA were higher in magnitude for foreign students than the local American group. The correlations between GRE scores and the GPA were however, weaker for foreign students than the respective median correlations on American students, computed from last several studies.

==========

Insert Table 3 Here

==========

The stepwise multiple regression was performed on foreign students to identify the best composite of multiple predictors and the accuracy with which these predictors can predict the GPA. The data was analysed in two different ways. The results of both the analysis were recorded in Table 4.

==========

Insert Table 4 Here

==========

The first analysis revealed that the GRE-V was the best single predictor of the GPA for foreign students among the predictors: GRE-Q, major, sex, year of enrollment, and their interactions. Major was found as the second best predictor and the only predictor that significantly improved the prediction of the criterion over and above the prediction

made by the GRE-V ($p < .01$). The remaining predictors and the interactions were not significant ($p > .05$). The highest multiple correlation of .46 was obtained when all the predictors and their relevant interactions were regressed on GPA.

In the second analysis, the regression of GRE-T did not lead to significant overall multiple correlation ($p > .05$). The value of R was lower than the overall R value obtained in the first analysis by regressing GRE-V and GRE-Q together. Major, again, significantly improved the prediction of GPA ($p < .01$). The rest of the variables were insignificant predictors of the criterion ($p > .05$). The maximum variance that could be explained by regressing all the predictors and the respective two-way interactions was 21.1% which was slightly lower than the value obtained in the first analysis.

Discussion

Compared to American students, lower means and higher standard deviations were the characteristics of foreign students, both on local and national level. High standard deviations on all the variables except GRE-V refer to the heterogeneity of the sample. It may be proposed thet foreign students be divided into more homogenious subgroups and that separate norms be established for each subgroup. Lower mean and lower standard deviation on the GRE-V portrayited language difficulty as a common problem of all the foreign students.

The cultural bias of the GRE Aptitude test for foreign students was evidenced by the t statistic. The results revealed that foreign students did not score significantly different from American students on the

GRE-Q, a scale which is less influenced by the language. However, foreign students scored significantly lower than Americans on the GRE-V which is heavily loaded with language component. Significantly lower mean GPA on foreign students may also be attributed to the verbal skills that dominate graduate studies and for which foreign students were not as well prepared as native speakers.

In spite of lower scores on the verbal section and higher potential for the verbal subscale to be culturally biased, the GRE-V turned out as the best single predictor of GPA for foreign students. It implied that verbal skills are comparatively more important than quantitative skills to succeed in the graduate school. The results of earlier validation studies, conducted on foreign students, were contradictory as to the significance of GRE-V and GRE-Q as predictors. Meberly (1960) found GRE-V as the best single predictor of GPA. Pitcher and Harvey (1963) found GRE-V as a significant predictor of GPA for the University of Florida, GRE-Q for the University of Texas, and GRE-V and GRE-Q both significant predictors at Florida State University and the University of Illinois. The finding of the present study about the significance of verbal scores therefore, could not be generalized neither across the population of foreign students in general nor to the local population of foreign students until similar results are found in subsequent replications.

The next best single predictor, though not statistically significant was the total score on the GRE (GRE-T). The utility of this predictor was however, limited due to its redundancy in the presence of subscale scores. The relationship of GRE scores with the GPA was, in

general, stronger for foreign students than the local group of American students. The lower r values on American students were the result of large sample size (n= 407) and thus were more stable and reliable than those of foreign students. This supposition was further supported when r values obtained on foreign students were compared with the respective median r values, reported by Conrad et al. (1977). It was observed that the relationshipp between GRE scores and the GPA is weaker for the group of foreign students than the average group of American students, used in the computation of median r values. Weak relationship between GRE scores and the GPA suggested that the GRE Aptitude test is not the most appropriate test to use for foreign students to predict their success in graduate school.

During multiple regression analysis, the GRE-V and major came out as the best composite of multiple predictors. This finding was consistent with the literature that reported GRE scores predicting differently for different majors. The general consensus is that the verbal scores are a better predictor of the GPA for descriptive majors and quantitative scores predict better for symbol oriented disciplines (Lannholm, 1972; Conrad et al., 1977). Another significant result of this analysis was the lower predictability of the total GRE scores compared to the independent use of subscale scores. Assigning equal weight to the verbal and quantitative scores was found undesirable.

Conclusion

This study was conducted to assess the differential predictive validity of the GRE Aptitude test. The results revealed that foreign

students scored significantly lower on GRE-V, GRE-T, and GPA compared to the local refrence group of American students. The scores on GRE-Q were found similar for native and non-native groups. The zero order correlations revealed GRE-V as the only statistically significant predictor of the GPA for foreign students. The magnitude of correlations between GRE scores and the GPA for foreign students was lower than the median correlation coefficients, computed from the GRE validation studies completed on American students. Major was found as the most significant predictor of GPA.

In the presence of low correlation between GRE scores and GPA, the GRE Aptitude test was not considered the most appropriate instrument in predicting academic success of foreign students. However, the significance of verbal scores, as a predictor of GPA, justified its continued use for non-natives. The establishment of separate norms for foreign students therefore, seemed highly desirable to minimize test bias caused by their consistently low performance on both the predictors and the criterion. High within group variability among foreign students suggested the study of subgroups formed on the basis of language, culture, or geograpphic location.

References

Angelis, P.J. Language testing and intelligence testing: Friends or Foes. Occasional Papers on Languages 1, 1977.

Conrad, L., Trismen, D., & Miller, R. (Eds.). Graduate Record Examinations Technical Manual. Princeton, N.J.: Educational Testing Service, 1977.

Harvey, P.R., & Lannholm, G.V. The performance of foreign students on the GRE Aptitude test (GRE Rep. 61-1). Princeton, N.J.: Educational Testing Service, 1961.

Kaiser, Javaid. The predictive validity of GRE Aptitude test. Paper presented at the meeting of Rockey Mountain Educational Research Association, Albuquerque, New Mexico, 1982.

Kerlinger, F.N., & Pedhazer, E.J. Multiple Regression in Behavioral Research. New York: Holt, Rinehart, and Winston, 1973.

Lannholm, G.V. Abstracts of selected studies of the relationship between scores on the GRE and graduate school performance (GRE Special Rep. 60-3). Princeton, N.J.: Educational Testing Service, 1960.

Lannholm, G.V. Summaries of GRE validity studies: 1966-1970 (GRE Rep. 72-1). Princeton, N.J.: Educational Testing Service, 1972.

Maberly, N.C. The validity of the GRE's as used with English speaking foreign students. Educational and Psychological Measurement, 1960, 23, 785-788.

Pitcher, B., & Harvey, P.R. The relationship of Graduate Record Examinations aptitude test scores and graduate school performance of foreign students at four American graduate schools (GRE Special Rep. 63-1). Princeton, N.J.: Educational Testing Service, 1963.

Shay, H.R., & Tseng, M.S. Language proficiency and aptitude as determinents of academic performance of foreign students in graduate school. Unpublished doctoral dissertation, West Virginia University, 1978.

Willingham, W.W. Validity and the Graduate Record Examinations. Princeton, N.J.: ETS, 1976.

TABLE 1

Sample size, Means, and Standard Deviations on Local and National Groups of American and Foreign Students

Variables	Local Foreign Students		Local American Students		National Sample of Foreign Students		National Sample of American Students	
	N = 148		N = 407		N = 21,881		N = 316,693	
	$\bar{X}$	S	$\bar{X}$	S	$\bar{X}$	S	$\bar{X}$	S
GRE-V	329.53	97.43	534.91	98.33	427	134	516	115
GRE-Q	513.11	173.53	525.13	123.37	523	142	528	132
GRE-T	842.63	215.31	1058.92	191.12	---	---	---	---
GPA	3.33	0.47	3.63	0.38	---	---	---	----

TABLE 2

Test of Means on the GRE Scores and the GPA between American and Foreign Students.

	Test for Homogeneity of Variance		Test for Equal Means	
Variables	df	F	df	t
GRE-V	1,553	1.02	553	21.81**
GRE-Q	1,204	1.96**	204	0.77
GRE-T	1,553	1.27	553	11.39**
GPA	1,218	1.59**	218	6.91**

*$P \leq .05$

**$P \leq .01$

TABLE 3

Intercorrelation between GRE Scores and GPA for American Students above the diagnol, for Foreign Students below the diagnol, and Median Correlations based on Previous Studies

Variables	GRE-V	GRE-Q	GRE-T	GPA
GRE-V	1.0	.46	.81	.14
GRE-Q	.20	1.0	.89	.04
GRE-T	.61	.90	1.0	.10
GPA	.18*	.08	.15	1.0
Median r	.24	.23	.33	---
No. of r's	46	43	30	---

*P $\leq$ .05

**P $\leq$ .01

TABLE 4

Test of Significance of R and Increment in R^2 by GRE Scores, major, sex, year of Enrollment and Their Interactions on Foreign Students

Variables	R	R^2	$\hat{R}^2$	R^2 Change	Test of Overall R df	Test of Overall R F	Test of Increment in R^2 df	Test of Increment in R^2 F
PART I								
GRE-V	.1789	.0320	.0254	.0320	1,146	4.826*	1,146	4.826*
GRE-Q	.1849	.0342	.0209	.0022	2,145	2.567	1,145	0.330
MAJOR	.3144	.0989	.0737	.0647	4,143	3.922**	2,143	5.134**
SEX	.3144	.0989	.0671	.0000	5,142	3.116*	1,142	0.000
YEAR	.3394	.1152	.0575	.0163	9,138	1.996*	4,138	0.636
MAJOR*SEX	.3784	.1432	.0739	.0280	11,136	2.066*	2,136	2.222
MAJOR*YEAR	.4398	.1934	.0737	.0502	19,128	1.616	8,128	0.996
SEX*YEAR	.4616	.2130	.0745	.0196	22,125	1.538	3,125	1.038
PART II								
GRE-T	.1467	.0215	.0148	.0215	1,146	3.212	1,146	3.212
MAJOR	.3144	.0989	.0801	.0774	3,144	5.266**	2,144	6.184**
SEX	.3144	.0989	.0737	.0000	4,143	3.922**	1,143	0.000
YEAR	.3393	.1152	.0642	.0163	8,139	2.261*	4,139	0.640
MAJOR*SEX	.3783	.1431	.0805	.0279	10,137	2.288*	2,137	2.230
MAJOR*YEAR	.4384	.1922	.0795	.0491	18,129	1.705*	8,129	0.980
SEX*YEAR	.4594	.2111	.0796	.0189	21,126	1.605	3,126	1.006

*$p \leq .05$

**$p \leq .01$